

周波数領域フィードフォワード形ブラインド信号源分離における信号歪みを抑制する学習法

著者	出島 康宏, 堀田 明秀, 中山 謙二, 平野 晃宏
雑誌名	電子情報通信学会技術研究報告. SIP, 信号処理 : IEICE technical report
巻	105
号	349
ページ	37-42
発行年	2005-10-01
URL	http://hdl.handle.net/2297/18184

周波数領域フィードフォワード形ブラインド信号源分離における 信号歪みを抑制する学習法

出島 康宏[†] 堀田 明秀[†] 中山 謙二^{††} 平野 晃宏^{††}

[†] 金沢大学大学院 自然科学研究科
920-1192 石川県金沢市角間町

E-mail: [†]{dejima,horita}@leo.ec.t.kanazawa-u.ac.jp, ^{††}{nakayama,hirano}@t.kanazawa-u.ac.jp

あらまし ブラインド信号源分離 (BSS) では、出力を独立にするように学習が行われるため信号歪みが発生する可能性があることが知られている。また、フィードフォワード形ブラインド信号源分離 (FF-BSS) は分離回路における自由度が高く、出力信号を互いに独立にする学習が信号歪みを生じる可能性がある。そこで信号歪みの低減のために完全分離の条件と無歪の条件から導かれた式を学習に加味するという手法がすでに提案されている。本稿では周波数領域における FF-BSS においてその手法を適用し、様々な信号源に対しての信号歪み抑制法の性能をシミュレーションにより解析する。

キーワード ブラインド信号源分離, 周波数領域, 信号歪, フィードフォワード, 学習

A Learning Algorithm Reducing Signal Distortion for Frequency Domain Feedforward Blind Source Separation

Yasuhiro DEJIMA[†], Akihide HORITA[†], Kenji NAKAYAMA^{††}, and Akihiro HIRANO^{††}

[†] Graduate School of Natural Science and Technology, Kanazawa Univ.
Kakuma-machi, Kanazawa, 920-1192

E-mail: [†]{dejima,horita}@leo.ec.t.kanazawa-u.ac.jp, ^{††}{nakayama,hirano}@t.kanazawa-u.ac.jp

Abstract In the blind source separation (BSS), a separation block is trained so as to make its output signals to be statistically independent. Since the frequency domain BSS has some degree of freedom, the wave form and frequency response of the separated signals may be changed from the originals, resulting in signal distortion. In this paper, a learning algorithm, suppressing the signal distortion, for the feedforward BSS, implemented in the frequency domain, is proposed. A condition on both complete separation and signal distortion free has been proposed. This condition is included in the learning process as constraint. Usefulness of the proposed learning algorithm is analyzed by using several kinds of the signal sources.

Key words Blind source separation, Frequency domain, Signal distortion, Feedforward, Learning

1. はじめに

ブラインド信号源分離においては、これまで数多くの方式が提案されてきている。混合方式では瞬時混合方式と畳み込み混合方式。学習領域では時間領域と周波数領域。回路構成としてはフィードフォワード (FF) 形、フィードバック (FB) 形といった様々な方式で、分離性能を向上させるための方法が提案されている [5]-[8], [11], [12], [14], [15]。ブラインド信号源分離の学習アルゴリズムとしては一般的に各出力の相互情報量を誤差関数とした自然勾配法が用いられている。しかし、自然勾配法は各出力を互いに独立にするように学習が進むため、出力された

信号が観測信号に含まれる信号源成分からずれてしまうことがある。これを信号歪みと呼ぶ [16]。特に、フィードフォワード形ブラインド信号源分離 (FF-BSS) は分離回路の自由度が高く、より信号歪みを発生しやすいといった問題がある。それに対し、完全分離、信号歪みなしといった条件から導かれる条件式を、従来の FF-BSS の学習アルゴリズムに付化することにより、信号歪みを抑制する学習法がすでに提案されている [17], [18]。

本稿では、周波数領域 FF-BSS に信号歪み抑制法を適用し、様々な信号源における抑制法の働きをシミュレーションにより解析する。

2. 畳み込み混合形 FF-BSS

2.1 回路構成

FF-BSS の簡単なものとして、信号源 2 つ、センサ数 2 つのモデルを用いる。ブロック図を図 1 に示す。

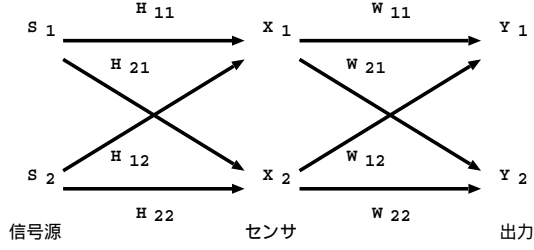


図 1: FF-BSS 回路構成

このとき、出力信号は以下の式で表される。

$$x_j(n) = \sum_{i=1}^2 \sum_{l=0}^{K_h-1} h_{ji}(l) s_i(n-l), j = 1, 2 \quad (1)$$

$$y_k(n) = \sum_{j=1}^2 \sum_{l=0}^{K_w-1} w_{kj}(l) x_j(n-l), k = 1, 2 \quad (2)$$

式 (1), (2) で表される $x_j(n)$, $y_k(n)$ は時間領域におけるものなので、これらを FFT を用いて周波数領域に変換すると、以下のように表される。

$$\mathbf{X}(e^{j\omega}) = \mathbf{H}(e^{j\omega}) \mathbf{S}(e^{j\omega}) \quad (3)$$

$$\begin{aligned} \mathbf{Y}(e^{j\omega}) &= \mathbf{W}(e^{j\omega}) \mathbf{X}(e^{j\omega}) \\ &= \mathbf{A}(e^{j\omega}) \mathbf{S}(e^{j\omega}) \end{aligned} \quad (4)$$

2.2 学習法

分離回路 $\mathbf{W}(r, m)$ の学習には、式 (2) で表される $y_k(n)$ を FFT を用いて周波数領域に変換した $\mathbf{Y}(r, m)$ を使用し、学習アルゴリズムは以下に示す出力の相互情報量をコスト関数とする自然勾配法を周波数領域に拡張したものをを用いる [4], [9], [10]。

$$\begin{aligned} \mathbf{W}(r+1, m) &= \mathbf{W}(r, m) \\ &+ \eta [\mathbf{I} - \langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle] \mathbf{W}(r, m) \quad (5) \\ \Phi(\mathbf{Y}(r, m)) &= \frac{1}{1 + e^{-\mathbf{Y}^R(r, m)}} + \frac{j}{1 + e^{-\mathbf{Y}^I(r, m)}} \quad (6) \end{aligned}$$

ここで、 η は学習係数、 r は学習回数、 m は m 番目の周波数帯域、 \mathbf{I} は単位行列、 \mathbf{Y}^R , \mathbf{Y}^I は出力 \mathbf{Y} の実部、虚部を表す。

また、式 (5) は出力を白色化する余分な制約があることが知られており、その問題を回避するために単位行列 \mathbf{I} を対角行列に置き換えた以下のような学習アルゴリズムがすでに提案されている [13]。

$$\begin{aligned} \mathbf{W}(r+1, m) &= \mathbf{W}(r, m) + \eta [\text{diag}(\langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle) \\ &- \langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle] \mathbf{W}(r, m) \quad (7) \end{aligned}$$

本稿では式 (5), (7) をそれぞれ学習法 I, 学習法 II と呼ぶことにする。

3. 信号歪み

3.1 信号歪み評価

信号歪みの基準として、信号源が考えられるが、BSS は混合過程を推定するものではなく、また、学習の規範は分離回路の出力信号における統計的な独立性であり、分離回路の出力信号を信号源に戻す作用はない。すなわち、BSS において、信号歪の基準を信号源にすることは、信号源分離の学習と整合しないことになる。そこで、本研究では、混合過程の伝達特性が推定できないという立場から、観測信号に含まれる信号源を信号歪みの基準とする。言い換えると、分離回路における信号波形や周波数特性の変化を「信号歪み」と見なすことにする。

出力信号 $Y_i(z)$ に含まれる信号源成分は $A_{ii}(z)S_i(z)$, $A_{ij}(z)S_j(z)$, ($i \neq j$)、観測信号 $X_i(z)$ に含まれる信号源成分は $H_{ii}(z)S_i(z)$, $H_{ij}(z)S_j(z)$, ($i \neq j$) によって表される。ここで、信号源 $S_i(z)$ が出力 $Y_i(z)$ に取り出されたと仮定すると、 $A_{ij}(z)S_j(z)$, $H_{ij}(z)S_j(z)$ は分離したい干渉音の成分となり、これらは信号分離で出力 $Y_i(z)$ 中から取り除かれるので、信号歪み評価としては、 $A_{ii}(z)S_i(z)$, $H_{ii}(z)S_i(z)$ の差を用いる。

さらに、信号に依存しない評価方法として伝達特性 $H_{ii}(z)$ と $A_{ii}(z)$ を比較する尺度も用いることにする。

$$\begin{aligned} \sigma_{d1a} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ii}(e^{j\omega})S_i(e^{j\omega}) \\ &- A_{ii}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \end{aligned} \quad (8)$$

$$\begin{aligned} \sigma_{d1b} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (|H_{ii}(e^{j\omega})S_i(e^{j\omega})| \\ &- |A_{ii}(e^{j\omega})S_i(e^{j\omega})|)^2 d\omega \end{aligned} \quad (9)$$

$$\sigma_1 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ii}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \quad (10)$$

$$SD_{1x} = 10 \log_{10} \frac{\sigma_{d1x}}{\sigma_1}, \quad x = a, b \quad (11)$$

$$\sigma_{d2a} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ii}(e^{j\omega}) - A_{ii}(e^{j\omega})|^2 d\omega \quad (12)$$

$$\sigma_{d2b} = \frac{1}{2\pi} \int_{-\pi}^{\pi} (|H_{ii}(e^{j\omega})| - |A_{ii}(e^{j\omega})|)^2 d\omega \quad (13)$$

$$\sigma_2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ii}(e^{j\omega})|^2 d\omega \quad (14)$$

$$SD_{2x} = 10 \log_{10} \frac{\sigma_{d2x}}{\sigma_2}, \quad x = a, b \quad (15)$$

3.2 重み付け効果

分離回路 $\mathbf{W}(r, m)$ の学習は式 (5), (7) の自然勾配法によって行なわれるが、その学習は出力 $\mathbf{Y}(r, m)$ を用いて行なわれる。分離回路 $\mathbf{W}(r, m)$ の初期値を $\mathbf{W}(0, m) = \mathbf{I}$ とした場合、出力は $\mathbf{Y}(0, m) = \mathbf{X}(0, m)$ となる。そして、 $\mathbf{X}(r, m) = \mathbf{H}\mathbf{S}(r, m)$ であるので、 $\mathbf{W}(r, m)$ の学習は $\mathbf{H}\mathbf{S}(r, m)$ に比例することが分かる。仮に、すべての信号源が音声であったとするならば、それらの周波数スペクトルは低域の周波数成分が高く、高域の周波数成分が低いという特徴を持つ互いに似た信号になる。そして、信号源が混ざりあってできる観測信号 $X_i(r, m)$ の周波数スペクトルも互いに似てくる。学習が進むにつれ出力 Y_k は徐々に信号源 S_i に近付いていくが、その際、分離回路 $\mathbf{W}(r, m)$ の学習はスペクトル毎の偏りを残したまま進む。すべての信号源

が音声のように似た周波数スペクトルを持つものである場合は、出力と信号源の周波数スペクトルに大きな差がないため分離に大きな問題はないが、信号源が音楽のように周波数スペクトルにばらつきがあり、それぞれが似ていない場合には、周波数成分の高い帯域は学習が進み、周波数成分の低い帯域は分離回路 $W(r, m)$ が初期値からあまり変化せず出力が信号源に近付かない可能性が考えられる。

このように周波数領域 FF-BSS では、重みを付けたように帯域毎に学習に偏りが生じる可能性があるため、信号源の性質によっては、そのことが信号分離を難しくし、且つ信号歪みを引き起こす原因になる場合があることが予想される。

4. 信号歪み抑制

本稿では簡単な FF-BSS として、図 1 で示した信号源 2 つ、センサ数 2 つの場合について考える。信号源 $S_i(z)$ が出力 $Y_i(z)$ に取り出された場合において、完全に信号分離がなされており、且つ信号が歪みなく出力されたとすると混合過程 $H_{ji}(z)$ と分離回路 $W_{ji}(z)$ の間に以下の関係式が成り立つ。

$$W_{11}(z)H_{11}(z) + W_{12}(z)H_{21}(z) = H_{11}(z) \quad (16)$$

$$W_{11}(z)H_{12}(z) + W_{12}(z)H_{22}(z) = 0 \quad (17)$$

$$W_{21}(z)H_{11}(z) + W_{22}(z)H_{21}(z) = 0 \quad (18)$$

$$W_{21}(z)H_{12}(z) + W_{22}(z)H_{22}(z) = H_{22}(z) \quad (19)$$

これらの式は 2 種類の意味を持つ。式 (17), (18) は信号源 $S_i(z)$ が出力 $Y_i(z)$, $i = 1, 2$ にそれぞれ取り出され、もう一方の信号源の成分が出力に含まれない状態。つまり、BSS によって完全に分離がなされた状態を表す。また、式 (16), (19) は観測信号 $X_i(z)$ に含まれる信号源成分 $H_{ii}(z)S_i(z)$ がそのまま出力 $Y_i(z)$ に取り出された状態。つまり、出力信号が歪みなしで取り出された状態を表す。

式 (17), (18) によって $H_{ji}(z)$ は以下の式の様に表される。

$$H_{12}(z) = -\frac{W_{12}(z)}{W_{11}(z)}H_{22}(z) \quad (20)$$

$$H_{21}(z) = -\frac{W_{21}(z)}{W_{22}(z)}H_{11}(z) \quad (21)$$

上式を式 (16), (19) に代入することで、 $H_{ji}(z)$ を消去すると以下の式が導かれる。

$$W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) = W_{22}(z) \quad (22)$$

$$W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) = W_{11}(z) \quad (23)$$

これらより、 $W_{11}(z) = W_{22}(z)$ が示されるので、式 (24) が成り立つ。

$$W_{jj}^2(z) - W_{jj}(z) - W_{jk}(z)W_{kj} = 0 \quad (24)$$

$$j = 1, 2, k = 1, 2, j \neq k$$

上式より、 $W_{11}(z), W_{22}(z)$ は以下の様に求められる。

$$W_{jj}(z) = \frac{1 \pm \sqrt{1 + 4W_{12}(z)W_{21}(z)}}{2}, j = 1, 2 \quad (25)$$

「完全分離」と「信号歪みなし」の両方の条件を満たすとき式 (25) が成り立つ。従って、式 (25) を式 (5), (7) の学習に加味することにより信号歪みの発生を抑制する。

$$W_{jj}(r, m) = (1 - \alpha)W_{jj}(r, m) + \alpha\tilde{W}_{jj}(r, m) \quad (26)$$

ここで、 W_{jj} は式 (5), (7) で更新される値、 \tilde{W}_{jj} は式 (25) で計算される値、 α はスケール係数である。 $\alpha = 0$ とすると信号歪みの式 (25) を加味しない形となり、 $\alpha = 1$ とすると、分離回路 W を式 (5), (7) で更新したのち、対角要素 $W_{ii}(z)$ を式 (25) で計算されたものに置き換える形になる。

5. シミュレーション

5.1 シミュレーション条件

FF-BSS はセンサ同士の距離が近くないと分離がうまくいかないことが知られている。よって、混合過程 $H_{ji}(z)$ は $H_{ji}(z) = 0.9z^{-1}H_{ii}(z)$ としてシミュレーションを行う。FFT 点数は 256 点、分離回路の初期値は $W_{11}(z) = W_{22}(z) = 1$, $W_{ji}(z) = 0, i \neq j$ とする。

信号分離性能評価としては、以下の 2 種類の SIR (信号対干渉信号比:Signal-to-Interference-Ratio) を評価関数として用いる。

$$\sigma_{s1} = \frac{1}{2\pi} \int_{-\pi}^{\pi} (|A_{11}(e^{j\omega})|^2 + |A_{22}(e^{j\omega})|^2) d\omega \quad (27)$$

$$\sigma_{i1} = \frac{1}{2\pi} \int_{-\pi}^{\pi} (|A_{12}(e^{j\omega})|^2 + |A_{21}(e^{j\omega})|^2) d\omega \quad (28)$$

$$SIR_1 = 10 \log_{10} \frac{\sigma_{s1}}{\sigma_{i1}} \quad (29)$$

$$\sigma_{s2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} (|A_{11}(e^{j\omega})S_1(e^{j\omega})|^2 + |A_{22}(e^{j\omega})S_2(e^{j\omega})|^2) d\omega \quad (30)$$

$$\sigma_{i2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} (|A_{12}(e^{j\omega})S_2(e^{j\omega})|^2 + |A_{21}(e^{j\omega})S_1(e^{j\omega})|^2) d\omega \quad (31)$$

$$SIR_2 = 10 \log_{10} \frac{\sigma_{s2}}{\sigma_{i2}} \quad (32)$$

5.2 異なる帯域を持つ有色信号

3.2 節で述べたように、周波数領域 FF-BSS では帯域の異なる信号源を用いることで学習の偏りが起こることが予想される。これは信号歪みを引き起こす原因となる。従って、信号源に異なる帯域を持つ 2 つの定常有色信号についてシミュレーションを行なった。観測信号 X と信号歪みの評価基準となる $H_{ii}(z)S_i(z)$ を図 (2), (3) に示す。

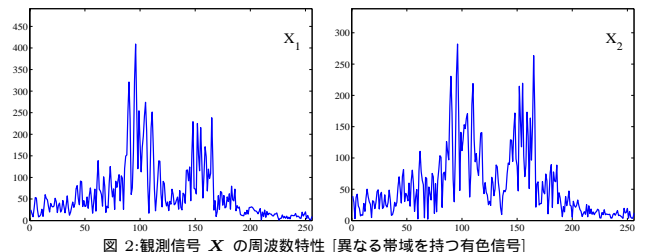


図 2: 観測信号 X の周波数特性 [異なる帯域を持つ有色信号]

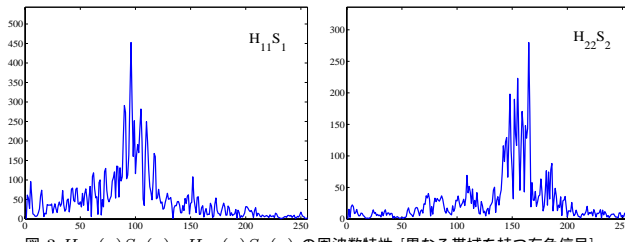


図 3: $H_{11}(z)S_1(z)$, $H_{22}(z)S_2(z)$ の周波数特性 [異なる帯域を持つ有色信号]

初めに, 学習法 I, II にて式 (25) の信号歪み抑制法を加味しない場合, つまり式 (26) にて $\alpha = 0$ として学習を行なった後の出力を以下に示す.

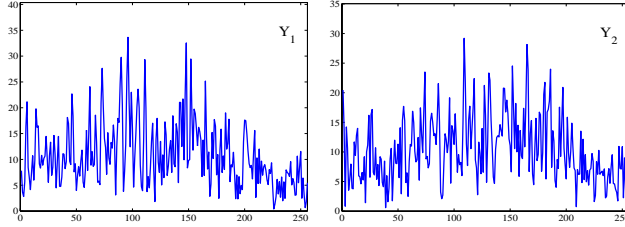


図 4: 学習法 I における出力 Y の周波数特性 [異なる帯域を持つ有色信号]

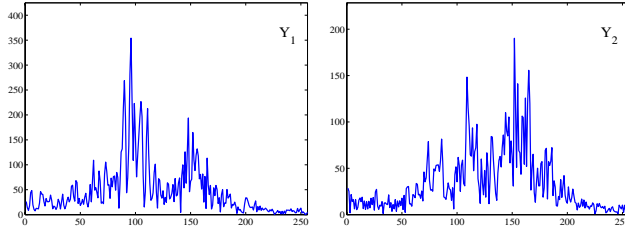


図 5: 学習法 II における出力 Y の周波数特性 [異なる帯域を持つ有色信号]

次に学習法 I, II にそれぞれ信号歪み抑制法を加味させていく. 式 (26) の α を学習法の学習係数 η に対して変化させていき, 信号分離性能と信号歪みの変化を見る. 学習係数は $\eta = 2 \times 10^{-6}$ として学習を行なった. 結果は以下の表 1, 2 のようになった.

表 1 学習法 I における信号分離, 信号歪み評価 ($\eta = 2 \times 10^{-6}$)
[異なる帯域を持つ有色信号]

	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
$\alpha = 0$	4.32	4.93	-3.61	-5.03	-3.67	-4.28
$\alpha = \eta$	3.29	3.68	-8.77	-10.01	-7.91	-9.05
$\alpha = \eta \times 10^3$	4.05	7.35	-17.62	-19.72	-10.87	-14.20
$\alpha = 1$	4.06	7.63	-18.36	-20.76	-10.81	-14.14

表 2 学習法 II における信号分離, 信号歪み評価 ($\eta = 2 \times 10^{-6}$)
[異なる帯域を持つ有色信号]

	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
$\alpha = 0$	2.82	5.22	-12.24	-14.34	-7.42	-9.44
$\alpha = \eta$	2.88	5.54	-12.92	-14.92	-7.79	-10.08
$\alpha = \eta \times 10^3$	2.53	5.32	-12.58	-14.73	-7.77	-10.59
$\alpha = 1$	2.72	5.67	-13.71	-15.66	-8.27	-11.07

図 4, 表 1 より, 学習法 I は信号歪み抑制法なしでは出力

が白色化し, 信号歪みが大きくなってしまふ. しかし, 式 (26) の α を徐々に 1 に近付けることによって信号歪みが大幅に低減されている. このことから, 抑制法が学習法 I の余分な制約から生まれる出力白色化を抑制し, 信号歪みを抑えながら分離が進むようにうまく働いていることが分かる. また, 学習法 II ではもともと出力は白色化しないので, 表 2 を見ても分かるように, $\alpha = 0$ の時も信号歪みが少なく分離が進んでいる. また, α を徐々に 1 に近付けていくと, 学習法 I と同様に信号歪みが低減されていることが見てとれる. しかし, 学習法 I に比べ学習法 II の方が抑制法の効果が小さい. これは, 帯域の異なる信号源を用いる場合, 学習法 II では 3.2 節で述べたマスキング効果が起こるため分離が難しいが, 学習法 I では学習に際して余計な制約により出力を白色化しようとするため, すべての帯域において学習が進む. そして, その後で抑制法により出力白色化を抑えるため, 重み付け効果の影響が学習法 II より少なく学習が進むためであると考えられる. $\alpha = 1$ とした時の出力を学習後の出力を以下に示す.

また, SIR_1 は学習法 I で $\alpha = 0$ のとき最も高い値が出ているが, これは信号を歪ませて値を稼いでいるためであるので, 信号歪みが抑えられている状態でしか分離評価に信頼性はないことがわかる.

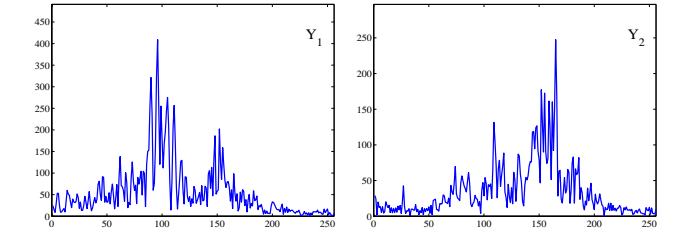


図 6: 学習法 I+抑制法における出力 Y の周波数特性 ($\alpha = 1$) [異なる帯域を持つ有色信号]

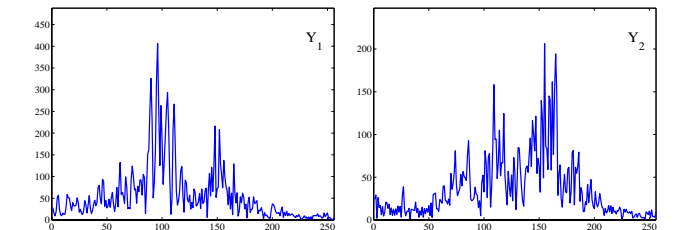


図 7: 学習法 II+抑制法における出力 Y の周波数特性 ($\alpha = 1$) [異なる帯域を持つ有色信号]

5.3 音声信号

次に信号源が帯域の似た信号の場合を考える. 2 つの信号源を互いに音声としてシミュレーションを行なった. $H_{11}(z)S_1(z)$, $H_{22}(z)S_2(z)$ の周波数スペクトルを以下に示す.

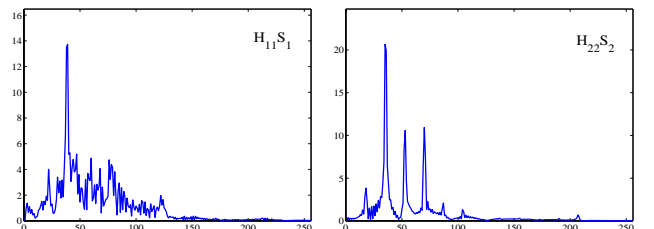


図 8: $H_{11}(z)S_1(z)$, $H_{22}(z)S_2(z)$ の周波数特性 [音声]

こちらも学習法 I, II にそれぞれ信号歪み抑制法を適用する.

表 3 学習法 I における信号分離, 信号歪み評価 ($\eta = 2 \times 10^{-6}$)

[音声]

	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
$\alpha = 0$	2.18	6.50	-6.73	-9.14	-5.38	-5.59
$\alpha = \eta$	2.60	6.91	-10.81	-13.72	-8.93	-9.67
$\alpha = \eta \times 10^3$	3.14	5.53	-10.90	-15.72	-12.31	-16.27
$\alpha = 1$	3.15	5.55	-10.90	-15.72	-12.31	-16.26

表 4 学習法 II における信号分離, 信号歪み評価 ($\eta = 2 \times 10^{-6}$)

[音声]

	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
$\alpha = 0$	2.53	5.83	-12.71	-17.79	-15.10	-19.34
$\alpha = \eta$	2.68	6.26	-13.92	-19.60	-16.34	-22.44
$\alpha = \eta \times 10^3$	2.77	6.49	-14.90	-20.85	-16.95	-23.55
$\alpha = 1$	2.77	6.49	-14.89	-20.85	-16.95	-23.55

信号源が音声の場合, 分離回路 W は信号の低域の部分の更新量が大きく, 高域の部分の更新量が小さくなる。つまり, 学習が低域に偏って進むが, 周波数成分がほとんどない高域は学習があまり行なわれなくても問題はなく, 低域の学習が進むことにより信号は分離される。従って, 帯域の異なる信号のように重み付け効果が学習を妨げることはない。

表 3, 4 より, 学習法 I, II ともに α が 1 に近くなるにつれ信号歪みが低減されていることが見てとれる。帯域の異なる信号源の場合とは違い, 学習法 I に比べ, 学習法 II の方が信号歪みは低減できているのは, 学習法 II において重み付け効果による学習の妨げがなく, むしろ重み付け効果が信号の周波数成分のないところを抑制している働きをしているためであると考えられる。つまり, 信号源の周波数特性が似ているものを分離する場合「学習法 II+抑制法」の方がより信号歪みなく分離できることが分かった。

また, 学習法 I, II 共に $\alpha = 1$ と $\alpha = \eta \times 10^3$ の信号歪みの値がほぼ同等の値を示したことから, 重み付け効果によって信号歪みが起こらない場合においては, α の値を小さくしても抑制法の効果があまり変わらないことが分かった。

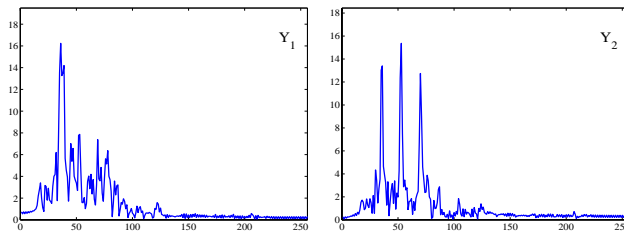


図 9: 学習法 I+抑制法における出力 Y の周波数特性 ($\alpha = 1$) [音声]

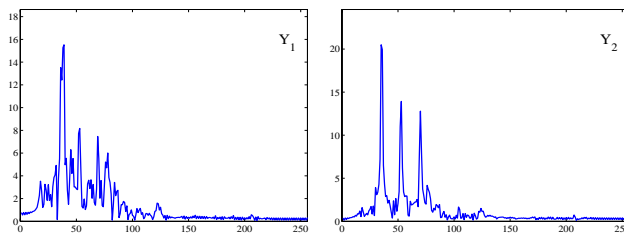


図 10: 学習法 II+抑制法における出力 Y の周波数特性 ($\alpha = 1$) [音声]

5.4 音 楽

これまで, 信号源の周波数成分が時間と共にほとんど変化しない場合についてシミュレーションを行ってきた。次に, 信号源に音楽を用いることによって, 信号源の周波数特性が時間と共に変化し, 且つ異なる帯域に周波数成分を持つ場合を考える。学習収束時の $H_{11}(z)S_1(z)$, $H_{22}(z)S_2(z)$ の周波数スペクトルは以下ようになった。

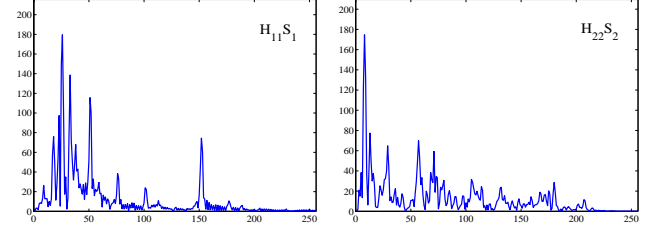


図 11: $H_{11}(z)S_1(z)$, $H_{22}(z)S_2(z)$ の周波数特性 [音楽]

シミュレーションの結果を以下に示す。

表 5 学習法 I における信号分離, 信号歪み評価 ($\eta = 2 \times 10^{-6}$)

[音楽]

	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
$\alpha = 0$	2.12	9.50	-4.31	-7.25	-1.63	-1.82
$\alpha = \eta$	2.80	7.80	-8.26	-11.65	-4.13	-4.49
$\alpha = \eta \times 10^3$	4.24	6.40	-12.62	-14.87	-10.13	-13.06
$\alpha = 1$	4.26	6.31	-12.57	-14.78	-10.20	-13.24

表 6 学習法 II における信号分離, 信号歪み評価 ($\eta = 2 \times 10^{-6}$)

[音楽]

	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
$\alpha = 0$	9.16	12.33	-9.39	-12.08	-8.28	-10.63
$\alpha = \eta$	7.70	10.35	-14.04	-16.66	-10.87	-13.48
$\alpha = \eta \times 10^3$	6.67	14.07	-18.59	-21.88	-11.48	-14.10
$\alpha = 1$	6.69	14.79	-18.90	-22.63	-11.56	-14.19

信号源が音楽の場合においては, 音声の場合と同様に「学習法 II+抑制法」がより信号歪みを低減できた。これは, 音楽が非定常信号であるため, 常に互いの信号源が異なる帯域の成分を持つとは限らず, 音声のように, 似たような帯域の成分を持つ時間が長かったためであると考えられる。

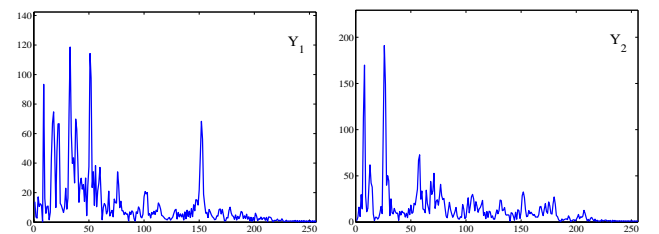


図 12: 学習法 I+抑制法における出力 Y の周波数特性 ($\alpha = 1$) [音楽]

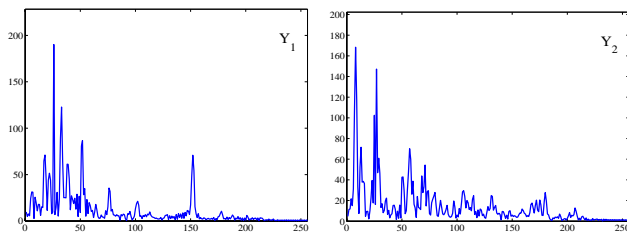


図 13: 学習法 II+抑制法における出力 Y の周波数特性 ($\alpha = 1$) [音楽]

6. ま と め

本稿では信号源の性質によって効果の大きさに違いはあっても、信号歪み抑制法が有効であることをシミュレーションにより確認した。また、信号源の性質によって起こりうる重み付け効果が信号歪みを生じる条件を確認し、周波数特性の似ている信号源を用いる場合は「学習法 II+抑制法」が、周波数特性が異なる信号源を用いる場合は「学習法 I+抑制法」が有効であることを示した。

文 献

- [1] C.Jutten and Jeanny Herault, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture," *Signal Processing*, 24, pp.1-10, 1991.
- [2] E.Sorouchyari, "Blind separation of sources, Part III: Stability analysis," *Signal Processing*, 24, pp.21-29, 1991.
- [3] H.L.Nguyen Thi and C.Jutten, "Blind source separation for convolutive mixtures," *Signal Processing*, vol.45, no.2, pp.209-229, March 1995.
- [4] S.Amari, T.Chen and A.Cichocki, "Stability analysis of learning algorithms for blind source separation," *Neural Networks*, vol.10, no.8, pp.1345-1351, 1997.
- [5] C.Simon, G.d'Urso, C.Vignat, Ph.Loubaton and C.Jutten, "On the convolutive mixture source separation by the decorrelation approach," *IEEE Proc. ICASSP'98*, Seattle, pp.IV-2109-2112, May 1998.
- [6] S.Cruces and L.Castedo, "A Gauss-Newton methods for blind source separation of convolutive mixtures," *IEEE Proc. ICASSP'98*, Seattle, pp.IV2093-2096, May 1998.
- [7] L.Parra and C.Spence, "Convolutive blind separation of nonstationary source," *IEEE Trans. Speech Audio Processing*, vol.8, pp.320-327, May 2000.
- [8] H.Mathis and S.C.Douglas, "On optimal and universal nonlinearities for blind signal separation," *IEEE Proc. ICASSP'01*, MULT-P3.3, May 2001.
- [9] S.Araki, S.Makino, T.Nishikawa and H.Saruwatari, "Fundamental limitation of frequency domain blind source separation for convolutive mixture of speech," *IEEE Proc. ICASSP'01*, Salt Lake City, MULT-P2.3, May 2001.
- [10] I.Kopriva, Z.Devcic and H.Szu, "An adaptive short-time frequency domain algorithm for blind separation of nonstationary convolved mixtures," *IEEE INNS Proc. IJCNN'01*, pp.424-429, July 2001.
- [11] K.Nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by kurtosis for blind source separation," *IEEE INNS, Proc. IJCNN'2002*, Honolulu, Hawaii, pp.1234-1239, May 2002.
- [12] K.Nakayama, A.Hirano and A.Horita, "A learning algorithm for convolutive blind source separation with transmission delay constraint," *IEEE INNS, Proc. IJCNN'2002*, Honolulu, Hawaii, pp.1287-1292, May 2002.
- [13] S.Araki, R.Mukai, S.Makino, T.Nishikawa, H.Saruwatari, "The fundamental limitation of frequency domain blind source separation for convolutive mixtures of speech," *IEEE Trans. Speech and Audio Processing*, vol.11, no.2, pp.109-116, March 2003.
- [14] K.Nakayama, A.Hirano and A.Horita, "A learning algorithm with adaptive exponential stepsize for blind source separation of convolutive mixtures with reverberations," *IEEE INNS, Proc. IJCNN'2003* July 2003.
- [15] K. Nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by estimated output pdf for blind source separation," *Proc. ICA2003*, Nara, Japan, pp.427-432, April 2003.
- [16] K.Nakayama, A.Hirano and Y.Dejima, "Analysis of signal separation and distortion in feedforward blind source separation for convolutive mixture," *Proc. MWSCAS2004*, Hiroshima, Japan, pp. III-207-III-210, July 2004.
- [17] 堀田明秀, 中山謙二, 平野晃宏, "フィードフォワード形およびフィードバック形ブラインド信号源分離における信号分離および信号歪の解析", *電子情報通信学会, 信学技法, SIP 研究会 (仙台)*, SIP2005-29, pp.25-30, Jun 2005.
- [18] A.Horita, K.Nakayama, A.Hirano and Y.Dejima, "Analysis of signal separation and signal distortion in feedforward and feedback blind source separation based on source spectra", *IEEE&INNS, Proc. IJCNN2005*, Montreal, pp.25-30, July-Aug 2005.